Amendments to the Specification:

In the Specification, please amend the following paragraphs as shown:

[0036] ~~FIG. 1 shows~~ FIGs. 1A-F show an embodiment of the invention for generating long DNA sequences from oligonucleotides immobilized on a surface, such as an oligonucleotide microarray.

[0037] ~~FIG. 2 shows~~ FIGs. 2A-G show an aspect of the invention for generating long DNA sequences from oligonucleotides synthesized on a surface, and then detached from that surface into solution.

[0038] ~~FIG. 3 shows~~ FIGs. 3A-D show an aspect of the invention for generating long DNA sequences starting from a set of many redundantly overlapped oligonucleotides, where the majority of the oligonucleotide sequence is used to generate the complementary overlap, thereby improving the possibility of annealing.

[0041] ~~FIG. 6 shows~~ FIGs. 6A and 6B show an embodiment of the invention employing an all-biological synthetic strategy for the synthesis of both single-stranded and double-stranded DNA using nucleotides with various 3'-phosphate protecting groups, such as but not limited to, peptide, carbohydrate, diphosphate, or phosphate derivative 3'-phosphate protecting groups.

[0043] ~~FIG. 8 shows~~ FIGs. 8A-C show an embodiment of the invention employing an all-biological synthetic strategy for the synthesis of double-stranded DNA using oligonucleotide hairpin-loops as heat-removable protecting groups.

[0044] ~~FIG. 9 shows on~~ FIGs. 9A and 9B show an embodiment of the invention employing force-feedback, in this case optical tweezers and/or a magnetic trap, in order to screen for and correct errors.

[0046] ~~FIG. 11 shows~~ FIGs. 11A and 11B show an embodiment of the invention employing parallel single-molecule systems using single and/or multiple arrays of light sources and

detectors to account for the possibility that an undetected and/or uncorrected error may have occurred and to ensure that the process is successful.

[0047] ~~FIG. 12 shows~~ FIGs. 12A and 12B show an aspect of the invention employing parallel single-molecule systems without arrays.

[0048] ~~FIG. 13 shows~~ FIGs. 13A and 13B show a method for the microfabrication of quadrupole arrays.

[0049] ~~FIG. 14 shows~~ FIGs. 14 A and 14B show an embodiment of the invention for error checking and error correction using nanopore devices for single-molecule synthesis with feedback using fluorescent 5' protecting groups.

[0050] ~~FIG. 15 illustrates~~ FIGs. 15A-G illustrate the independent control of a cluster of superparamagnetic beads by an electric field and opposing magnetic field gradient.

[0051] ~~FIG. 16 shows~~ FIGs. 16A-C show an embodiment of the invention for removing error sequences using mismatch binding proteins (MMBP). An error in a single strand of DNA causes a mismatch in a DNA duplex, which is selectively bound by a MMBP and separated from error-free DNA by methods known in the art such as by affinity capture or mobility differences.

[0055] ~~FIG. 20 shows~~ FIGs. 20A and 20B show an aspect of the invention for removing and correcting error sequences using a MMBP fusion to a non-specific nuclease. The MMBP binds to a mismatch in a DNA duplex; the MMBP-bound DNA complex is then removed using methods of protein purification.

[0056] ~~FIG. 21 shows~~ FIGs. 21A and 21B show an aspect of the invention for removing and correcting error sequences using a MMBP fusion to a non-specific nuclease and both strand invasion and branch migration to synthesize the error-free portions of the nucleotide sequence.

[0058] ~~FIG. 23 shows~~ FIGs. 23A and B show an aspect of the invention for removing and correcting error sequences using a non-specific endonuclease to cut the molecule into shorter

strands, binding error containing strands with MMBP, separating error containing MMBP complexed strands, and annealing and ligating cohesive ends.

[0064] ~~Figure 1 displays~~ FIGs.1A-F display a process of the invention for generating long DNA sequences from oligonucleotides immobilized on a surface, such as an oligonucleotide microarray. Such arrays are currently generated by a variety of synthetic approaches, including photolabile deprotection, photo-induced acid-labile deprotection, electrically-induced acid-labile deprotection, and inkjet printing of reagents. The number of different oligonucleotides that can be produced in microarray form is quite large. Some arrays may hold about 20,000 distinct locations, each with a different oligonucleotide sequence. The highest density arrays can contain about 400,000 distinct locations per square centimeter. For an array of 50-mer oligonucleotides, this would correspond to 20 million bases, roughly four times the genome size of many common bacteria.

[0071] One advantage of the method shown in ~~Figure 1~~ FIGs. 1A-F is the ability to track the progress and growth of the product by fluorescence. The free oligonucleotide corresponding to the 5' end of the sequence can include a fluorescent group at the 5' terminus. As the growing chain anneals to different spots on the microarray, regions of high concentration of the fluorescent group (where the free oligonucleotide is bound) are detected by fluorescence microscopy. Thus, the progress of the growing chain can be monitored. For example, fluorescence at oligonucleotide spot 3 indicates that the growing free DNA chain must contain at least sequence ABC in order to anneal. This monitoring is especially useful in the case of potential mis-annealing between sequences which are similar, but not the intended (perfect) match. In this case, the presence of a fluorescent spot at an unexpected location shows which sequence the free oligonucleotide has annealed to.

[0072] Another aspect of this invention is the stepwise repositioning of the growing DNA chain as a means to control the movement of some additional component. Referring to FIGs. 1A-F ~~figure 1~~, for the first cycle of annealing, the attached component will only be present at spot 1. Following strand extension by polymerase, the sequence attached to the component now has the sequence AB. In the second cycle of annealing, this complex will advance no further than spot 2, and so forth. The attached growing chain will still also have affinity for spot 1, and will be

partially localized there as well. However, the component and attached DNA chain can be "chased" through the spot locations by adding an excess of free oligonucleotide sequence A in a later cycles. For example, adding excess A in cycle 2 means that free oligonucleotide A will compete with the AB-attached component to anneal to spot 1 (in essence, flushing the AB-attached component away from this site), but only the AB-attached component will have affinity for site 2 (via the interaction between B-B' sequences).

[0074] ~~Figure 2 shows~~ FIGs. 2A-G show a process of the invention for generating long DNA sequences from oligonucleotides synthesized on a surface, and then detached from that surface.

[0084] ~~Figure 3 shows~~ FIGs. 3A-D show the synthesis of a large DNA molecule starting from a set of many redundantly overlapped oligonucleotides. As in FIGs. 2A-G ~~Figure 2~~, assembly relies on annealing complementary pairs of oligonucleotides and extending them to produce longer segments of DNA, until the full-length sequence is produced. However, in this case, the majority of the oligonucleotide sequence is used to generate the complementary overlap, improving the maximum possible specificity of annealing. Though the first polymerase extensions only produce slightly larger pieces of DNA, later growth steps are still exponential. Also, sometimes a particular oligonucleotide synthesis may fail, or be especially inefficient. For methods and materials known in the art related to nucleotide synthesis involving overlapped oligonucleotides see, e.g., European Patent Application EP 1314783A1 titled *Nukleinsaure-Linker und deren Verwendung in der Gensynthese* assigned to Sloning BioTechnology GmbH, which is hereby incorporated by reference.

[0085] This approach provides "insurance" against the failure of the synthesis of any one distinct oligonucleotide sequence. For example, in FIGs. 2A-G ~~Figure 2~~, a failure to produce oligonucleotide sequence CD would result in an inability to produce the longer CDE and ABCDE strands. In contrast, removal of any one oligonucleotide shown in Figure 3B does not prevent assembly of the full-length molecule. Thus, the many possible overlaps ensure that even if one oligonucleotide (such as oligonucleotide 2) were removed, the full-length assembly will still be achievable, because the full-length sequence is encoded redundantly in multiple oligonucleotides.

[0101] ~~Figure 8.~~ FIGs. 8A-C. Synthesis of double-stranded DNA using oligonucleotide hairpin-loops as heat-removable protecting groups. Oligonucleotides with secondary conformational structures, such as DNA hairpin-loops (also termed stem-loops, and molecular beacons), can also be used as protecting groups. A similar approach has been reported where hairpin-loops are enzymatically removed by restriction enzymes, a process termed "synthetic cloning" or "splinking." The methods described in FIGs. 8A-C ~~figure 8~~ differ from previously reported methods in the general structure of the hairpin-loops, and because the removal method is gentle heating. Furthermore, gentle heating is potentially an improved method of deprotection over enzymatic removal because 1) heat distributes more quickly and uniformly than enzymes because the enzymatic removal rate is diffusion-limited, and 2) gentle heating is a lower-cost resource than restriction enzymes.

[0118] ~~Figure 9~~ FIGs. 9A and 9B. Force-feedback using magnetic and optical tweezers.

[0123] Heat may also be used as an additional feedback and error correction mechanism in force feedback systems. For example, the force-feedback systems shown in FIGs. 9A-B ~~Figures 9~~ and 10 can also employ heat as additional feedback and error-correction. Prior to enzymatic ligation, the melting point of the small oligonucleotide in contact with the growing nucleic acid strand will be lowered if base-pair mismatches occur. The controlled application of heat after detected annealing can provide additional feedback about base-pair mismatches. If the oligonucleotide dehybridizes from the growing strand as the melting point is approached, but not reached, a base-pair mismatch is detected when a decrease in magnetophoretic force, or increase in electrophoretic force is required to keep the bead in equilibrium. Because the erroneous strand is removed by heat, this feedback process is also an error-correction mechanism.

[0124] Nucleotide removal by exonuclease activity may also be used for error-correction in force-feedback systems. The schemes in the force-feedback systems shown in FIGs. 9A-B ~~Figures 9~~ and 10 may also employ nucleotide removal by exonuclease activity as an error-correction mechanism. This type of error-correction is particular useful for correcting errors after enzymatic ligation of an erroneous strand. Whereas it would be extremely difficult to control the exact number of nucleotides that exonuclease removes from the 3'-end of a growing strand of nucleic acid, that level of control is not required in the methods reported herein because

the feedback systems allow for the length of the strand to be determined after the error-correction steps. Therefore, if too many nucleotides are initially removed, they may be added back later.

[0134] ~~Figure 14.~~ FIGs. 14A and 14B. Nanopore devices for single-molecule synthesis.

[0136] Figure 14B shows the design of a second nanopore device for single-molecule synthesis with feedback using fluorescent 5' protecting groups. Monitoring the deprotection of the 5' group is necessary to eliminate deletion errors. In this device, the growing strand is deprotected, and the wash is flowed through the nanopore, not the bead, and the nanopore only leads to one channel. If no fluorescence is detected in the wash, then the strand was not deprotected, or it was successfully deprotected but the fluorescent protecting group was not detected. The wash is constantly recycled until a fluorescent group is detected. Because there are no free nucleotides (only the growing strand) in this device, no addition error can occur by redundant 5' deprotection steps. Once the freed protecting group is detected, the bead is passed back to the device described in Figure 9A for a subsequent base addition. ~~For~~ Many methods and materials are known in the art relating to nanopore analysis~~.~~.

[0137] ~~Figure 15 shows~~ FIGs. 15A-G show an example of the independent control of a cluster of superparamagnetic beads by an electric field and opposing magnetic field gradient. These are screenshots obtained from a CCD camera mounted on a microscope. In each screenshot, the electrophoretic force moves the beads to the left of the screen, and the magnetic field gradient moves the bead to the right of the screen (i.e. the positive electrode is outside and towards the left of the field-of-view, and the magnetic tweezer apparatus is outside and towards the right of the field-of-view).

[0143] ~~Figure~~ FIGs. 15F and 15G. The experimental system schematic is shown in ~~figure~~ FIGs. 15F and 15G below, and the experimental details can be found in the accompanying description of FIGs. ~~figure~~ 15F and 15G. ~~Figure~~ FIGs. 15F and 15G depict ~~depicts~~ a method for the construction of an electrophoretic reservoir and magnetic tweezer. Superparamagnetic beads 1.05μm in diameter were obtained from Dynal Biotech (DynaBeads MyOne Carboxylic Acid). Beads were washed according to standard protocols and dispersed in distilled water. The electrode structure was made by thermal evaporation of aluminum on a glass slide. The

electrodes were spaced apart by about 1 cm using kapton tape as a mask. The reservoir was created by first placing an o-ring between the aluminum pads, and then sealing the reservoir with a glass cover slip. The single-pole magnetic tweezer was placed approximately 3 mm from the ground electrode, such that the attractive magnetic field gradient opposed the electrophoretic force felt by the beads. The single-pole magnetic tweezer was composed of a tip-pole electromagnet with a laser-cut scaffold to bring the tip of the tweezer as close to the top coverslip as possible. The core of the electromagnet was about 25 mm in length and about 10 mm in diameter. It was wrapped about 300 times with insulated copper wire that was potted using epoxy. The tips of the electromagnets were cut at about a 45° using a diamond saw. The current through the electromagnet and voltage across the electrodes were controlled using custom written software written in Labview. The entire apparatus was placed on the stage of a custom built optical microscope with a 20x condenser lens and 100x objective lens. Images were collected using a CCD camera and frame grabber that output to the software.

[0145] ~~Figure 16.~~ FIGs. 16A-C. Removal of error sequences using mismatch binding proteins. An error in a single strand of DNA causes a mismatch in a DNA duplex. A mismatch recognition protein (MMBP), such as a dimer of MutS, binds to this site on the DNA.

[0151] Many biological DNA repair mechanisms rely on recognizing the site of a mutation (error) and then using a template strand (most likely error-free) to replace the incorrect sequence. In the de novo production of DNA sequences, this process is complicated by the difficulty of determining which strand contains the error and which should be used as the template. In this invention, the solutions to this problem rely on using the pool of other sequences in the mixture to provide the template for correction. These methods can be very robust: even if every strand of DNA contains one or more errors, as long as the majority of strands have the correct sequence at each position (expected because the positions of errors are generally not correlated between strands), there is a high likelihood that a given error will be replaced with the correct sequence. ~~Figures 18-24~~ FIGs. 18, 19, 20A-B, 21A-B, 22, 23A-B, and 24 present procedures for performing this sort of local error correction.

[0154] In a preferred embodiment of the invention, errors are detectable in the form of a DNA mismatch, and can be removed by the combined action of 1) a protein, molecule, or process

which recognizes mismatches; and 2) a second protein, molecucle, or process which cleaves the DNA. Figure 19 illustrates a process for removing errors utilizing a mismatch recognition function in cooperation with a DNA cleavage agent. ~~Figure 20 demonstrates~~ FIGs. 20A and 20B demonstrate one possible design for an agent capable of combining these two functions.

[0157] ~~Figure 20.~~ FIGs. 20A and 20B. A protein designed to combine the functions of error recognition and error removal. The gene for a mismatch recognition protein (such as MutS) has been linked to the gene for a nuclease domain (such as that of restriction endonuclease FokI). when this gene is expressed, both functions will be combined in the same protein molecule, which will contain two separately folded domains. As MutS forms a dimer, so will this designed protein, allowing it to bind DNA at the site of a mismatch and cut both strands of DNA, excising the segment which contains an error, as shown in Figure 19. In a preferred embodiment of the invention, the designed protein would be thermostable. For example the binding and nuclease domains could be derived from thermophilic organisms, or proteins could be engineered for thermostability. This feature would allow the protein to function in a thermally cycled reaction, such as PCR or LCR, allowing error correction to occur in tandem with assembly of molecules of nucleic acid.

[0159] Figure 20B. A single tube process for assembling or amplifying molecules of nucleic acid while correcting errors. A tube or chamber for thermocycled reactions is divided into two regions, separated by a membrane. As the nucleic acids are assembled (or amplified), a thermostable protein (as in Figure 20A) acts on the nucleic acid to remove errors. The small pieces of excised error-containing DNA are the only ones small enough to pass through the membrane to the other side of the chamber. Here they encounter a resin with affinity for nucleic acid, so that they are not able to pass back into the other chamber, and are effectively removed from the desired nucleic acid product. Reassembly of the nucleic acid molecules surviving this process can be accomplished in many ways (see Figures 19, 21A-B, ~~and~~ 22, and 23A-B), including a PCR reaction which can take place in the same reaction. Multiple thermal cycles dissociate and reassociate the DNA duplexes. Where errors may still be present, this reassortment of individual strands provides new templates for error correction.

[0160] ~~Figure 21 follows~~ FIGs. 21A and 21B follow a process similar to that of Figure 19. However, in this embodiment of the invention, double-stranded gaps in DNA duplexes are repaired using the protein components of a recombination repair pathway. (Note that in this case global melting and re-annealing of DNA strands is not an absolute requirement , which can be preferable when dealing with especially large DNA molecules, such as genome length DNA.)

[0163] It is important to make clear that the methods of this invention are capable of generating large error-free DNA sequences, even if none of the initial DNA products are error-free. Figure 22 summarizes the effects of the methods of Figure 19 (or equivalently, ~~Figure~~ FIGs. 21A-B) applied to two DNA duplexes, each containing a single base (mismatch) error.

[0164] Figure 22. Two DNA duplexes are shown, identical except for a single base mismatch in each, at different locations in the DNA sequence. Mismatch binding and localized nuclease activity are used to generated double-stranded breaks which excise the errors. Recombination repair (as in FIGs. ~~Figure~~ 21A-B) or melting and reassembly (as in Figure 19) are employed to generate DNA duplexes where each excised error sequence has been replaced with newly synthesized sequence, each using the other DNA duplex as template (and unlikely to have an error in that same location). Note that complete dissociation and re-annealing of the DNA duplexes is not necessary to generate the error-free products (if the methods shown in FIGs. ~~Figure~~ 21A-B are employed).

[0165] A simple way to reduce errors in long DNA molecules is to cleave both strands of the DNA backbone at multiple sites, such as with a site-specific endonuclease which generates short single stranded overhangs at the cleavage site. Of the resulting segments, some are expected to contain mismatches. These can be removed by the action and subsequent removal of a mismatch binding protein, as described in Figure 19. The remaining pool of segments can be re-ligated into full length sequences. As with the approach of FIGs. ~~Figure~~ 21A-B, this approach includes several advantages. 1) loss of an entire full length DNA duplex is not required to remove an error; 2) global dissociation and re-annealing of DNA duplexes is not necessary; 3) error-free DNA molecules can be constructed from a starting pool in which no one member is an error-free DNA molecule.

[0168] ~~Figure 23 illustrates~~ FIGs. 23A and 23B illustrate the semi-selective removal of mismatch-containing segments.

[0170] Figure 23B. Fragments bound to MMBP are removed from the pool, as described in FIGs. 6A and 6B ~~Figure 6~~. The cohesive ends of each fragment allow each DNA duplex to associate with the correct sequence-specific neighbor fragment. A ligase (such T4 DNA ligase) is employed to join the cohesive ends, producing full length DNA sequences. These DNA sequences can be error-free in spite of the fact that none of the original DNA duplexes was error-free. Incomplete ligation may leave some sequences which are less than full-length, which can be purified away on the basis of size.